

Individual variations in infectiousness explain long-term disease persistence in wildlife populations

Stephanie Kramer-Schadt, Néstor Fernández, Dirk Eisinger, Volker Grimm and Hans-Hermann Thulke

S. Kramer-Schadt (stephanie.kramer-schadt@bio.uib.no), N. Fernández, D. Eisinger, V. Grimm, and H.-H. Thulke, UFZ, Helmholtz Centre for Environmental Research – UFZ, Dept of Ecological Modelling (OESA), Permoser Str. 15, DE-04318 Leipzig, Germany. SKS also at: Dept of Biology, Univ. of Bergen, Thormøhlensgate 55, NO-5008 Bergen, Norway. NF also at: Dept of Applied Biology, Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas, Avenida Maria Luisa s/n, ES-41013 Sevilla, Spain.

Viral disease persistence in species without a reservoir host is of importance for public health and disease management. But how can disease persistence be explained? We developed a spatially-explicit individual-based model that takes into account both ecological and viral traits as well as variable space to test disease persistence hypotheses under debate. We introduce a novel concept of modeling alternative disease courses at the individual level, causing transient infections or killing infected animals, with the lethally infected having a variable life-expectancy. We systematically distinguish between disease invasion and persistence. We use classical swine fever (CSF), an economically very important livestock disease in a social host, the wild boar, as a reference system to test and rank the persistence hypotheses under debate. Parameter values for host population demographics and CSF epidemiology reflect current knowledge. Sensitivity analysis of the model parameters revealed that the most important factor for disease persistence is a disease profile with mostly transient, i.e. surviving individuals requiring immunity, and some chronically, long-term infected animals. Immune individuals can constantly produce susceptible offspring, while some chronically infected individuals act as ‘super spreaders’ in time. Thus, variations in the course of the disease at the individual level are important factors determining persistence, which is usually not taken into account in the prominent measure of epidemiology, i.e. the basic reproductive number R_0 , which reflects the ‘reproductive potential’ of the infected sub-population. We discuss our results with regard to the general issues of modeling epidemics and disease management issues.

A major issue for controlling diseases is to understand how viral pathogens can persist within their host’s population without a reservoir (Caley and Hone 2006). Three major factors may influence the long-term dynamics of infectious diseases in the population: (1) the properties of the virus, (2) life history traits and population parameters of the host, and (3) the spatial structure of the host population. Several viral properties cause long-term infectiousness, for example, by remaining dormant in once-infected hosts and, upon reactivation, causing hosts to become infectious (Mollema et al. 2005). In addition, several host traits play important roles in driving ongoing disease dynamics. Breeding seasonality can force recurrent epidemics through the input of fresh susceptibles (Lloyd-Smith et al. 2005a, Conlan and Grenfell 2007). Seasonal social aggregation, for example during reproduction, influences transmission dynamics and therefore the spread of diseases (Hosseini et al. 2004), while parameters such as social group size, recruitment rate and movement can affect persistence (Cross et al. 2005, 2007, Rossi et al. 2005b). Third, the spatial distribution of the host population can have an effect on persistence in several ways. For instance, locally unstable host–virus interactions

may persist because of spatial separation in homogeneous environments because different areas are in different dynamic phases (Hagenaars et al. 2004). Moreover, there is often spatial variability in local factors affecting population parameters (reviewed by Orive et al. 2005).

Remarkably, the traditional susceptible-infected-recovered (SIR) mean-field models have predominated in the epidemiological literature for explaining disease outcomes, but their simplicity limits the exploration of more complex hypotheses involving the interacting effects of different pathogen, host and spatial parameters (Hansen et al. 2004). Thus, there is a need to take into account ecological factors and to incorporate ecological models in epidemiology for developing protection strategies (Kelly et al. 2003, Gewin 2004, Mackinnon and Read 2004, Eisinger and Thulke 2008). Moreover, recent evidence gleaned by modeling the extinction risk of small populations suggests that individual variability can be decisive for the outcome of population dynamic processes. Even a small number of individuals that are less affected by environmental variation can buffer population persistence (Grimm et al. 2005). The impact of particular individuals on epidemics has been conceptualized

under the term ‘super spreaders’ (Galvani and May 2005, Lloyd-Smith et al. 2005b), that is, individuals with a disproportionate effect on the course of the disease.

However, the effect of individual variability on disease persistence is still undervalued in epidemiological studies. Therefore, we developed a simulation model to disentangle the relative contributions to disease persistence of a variety of host and viral traits as well as the spatial effect of the host distribution. We define persistence as an endemic, recurrent infection within a closed, spatially restricted population. Thus, we distinguish disease persistence, the endemic phase, from disease invasion, the epidemic phase (Lloyd-Smith et al. 2005a, Conlan and Grenfell 2007), and then analyze what factors lead to long-term persistence after successful invasion. To evaluate the reliability of analytical mean-field measures for assessing persistence in more complex scenarios, we introduce \hat{R}_t , an individual-based measure related to the epidemiological standard value R_0 and containing elements of the effective reproduction number R_t (Anderson et al. 2004). The basic reproductive number R_0 is the expected number of secondary infections caused by an infective agent in a completely susceptible population (Anderson and May 1991). We investigate how well the analytical threshold $R_0 > 1$ performs relative to \hat{R}_t under heterogeneous conditions to predict persistence.

We use classical swine fever (CSF) in a social host, the wild boar *Sus scrofa*, as a reference system because it is of economical and ethical concern. Recent research has concentrated on aspects of disease eradication (Bieber and Ruf 2005, Choisy and Rohani 2006, Bolzoni et al. 2007). However, so far it is unclear how the disease can persist in some spatially limited foci (Fritzemeier et al. 2000). There are at least five hypotheses explaining persistence (Table 1),

including high local densities or large distribution areas of the host, the evolution of the virus towards moderate virulence, or prenatally infected offspring causing long-term infectiousness (reviewed by Kramer-Schadt et al. 2007). There is even a debate on whether the disease can persist in wild boar at all, or whether new outbreaks are due to repeated virus introductions by humans. An understanding of the causes leading to persistence is the prerequisite for launching appropriate eradication measures.

Our model is the first one to explore persistence of CSF in wild boar, where little is known about the processes in the field, underlining the timeliness of the research. The high diversity and complexity of the factors that potentially influence CSF persistence limits the utility of traditional SIR models for understanding the disease dynamics, making a more complex model necessary to assess all these factors simultaneously. We use the model to contrast and rank the five different hypotheses currently under debate (Table 1), involving viral, host and spatial factors. Finally, we discuss how individual-based simulation models can be used for addressing hypotheses involving ecological–epidemiological systems and disease management issues.

Material and methods

The model

We describe the model following the ODD protocol (overview, design concepts, details in Grimm and Railsback 2005, Grimm et al. 2006). The last element, containing the detailed description of the sub-models as well as model parameters, forms the Supplementary material Appendix 1.

Table 1. Hypotheses of disease persistence related to classical swine fever (reviewed by Kramer-Schadt et al. 2007).

Hypothesis	Description
1 – Density (number of individuals per space unit, in the model ‘carrying capacity per cell’) (Guberti et al. 1998, McCallum et al. 2001)	Transmission is seen as a density-dependent process. Above a certain threshold of susceptible individuals, diseases are likely to persist. The widespread increase in the distribution and size of wild boar populations in Europe since the early 1950s may have favored CSF persistence.
2 – Space (spatial extent, landscape size or distribution area, in the model ‘number of habitat cells’) (McCallum et al. 2001, Artois et al. 2002)	Works in non-spatial models in the same way as density (via number of animals). In spatial models, it brings a time delay until the primary invasive outbreak has run through the whole population, so that newborn susceptibles of immune survivors can be infected (frequency-dependent transmission), therefore contributing to disease persistence.
3 – Long-term virus shedders (Kern et al. 1999)	Prenatally infected piglets develop persistent viraemia and can survive for a long time maintaining the disease in the population, but eventually all die.
4 – Moderate virulence (Meyers and Thiel 1996)	The outcome of a CSF infection at the individual level, i.e. being transient (individuals recover after being infectious for 1–2 weeks; leads to immunity), acute (lasts less than 1 month; lethal) or chronic (lasts longer than 1 month; lethal), is related to the virulence of the virus, the age of the pig and its immune response (Dahle and Liess 1992). The involvement of viral strains of moderate to low virulence that are less pathogenic (transient infections) and that cause more chronic infections with prolonged viraemia is thought to be a factor for disease persistence.
5 – Disease spread through dispersal (Depner et al. 2000)	If the level of maternal antibodies in piglets and sub-adults is still low at the time of infection, i.e. shortly after they have lost full protection, these animals become transiently infected. It is presumed that the transiently infected animals are still able to disperse, spread the virus and thus contribute to disease persistence.

State variables and scales

The model consists of two sub-models, a demographic model of wild boar considering seasonal reproduction, natal dispersal and mortality, and a CSF virus model that is based on interventions in these ecological processes. The state variables of the wild boar individuals are (1) sex, (2) age in weeks (referred to as age class), leading to classification of piglets (**B**8 months), sub-adults (**B**2 years) and adults, (3) family group ID, (4) location (x-y coordinates), (5) demographic status (disperser or resident), and (6) epidemiological status, i.e. susceptible, immune, transiently infected, or lethally infected with a variable infectious period depending on an acute or chronic disease course.

The landscape consists of a grid of square cells that represent wild boar home ranges (Fernandez et al. 2006) and was of variable extent. A home range is assumed to represent 4 km² (Leaper et al. 1999). Cells represent home ranges of the same quality regarding density, composing a homogeneous habitat patch. Time steps represent one week as this is approximately the incubation period of the disease (Artois et al. 2002). Simulations lasted for 10 years or until the disease went extinct.

Process overview and scheduling

Reproduction, female natal dispersal and mortality are calculated in discrete weekly time steps (see Supplementary material Appendix 1). The density-dependent reproduction of the host is characterized by the local breeding capacity per cell or home range, i.e. the maximum number of females that can have offspring, and a seasonal reproduction pattern. Wild boar groups are linked by neighbourhood contact (eight surrounding cells) and female dispersing groups. Dispersal is characterized by a maximum number of movement steps per week. Male dispersal is neglected here because females establish new family groups (Dardallion 1988).

At the beginning of each year, yearly parameters like survival rates of the different age classes are assigned, because these are stochastic events that resemble good or bad years for boars. At the beginning of the weekly time step, infection takes place, then females might disperse and reproduce, then wild boar die due to infection or baseline mortality probability. At the end of each weekly time step, the age and duration of the infection in the individual are updated.

Design concepts

The model considers wild boar demography, infection within and between groups, natal dispersal, and interaction of the virus with these ecological processes (Supplementary material Appendix 1). The behavior and demography of the wild boar are imposed by age-dependent parameters. Stochasticity is included to represent demographic and environmental noise. The virus sub-model is also probabilistic. It includes a binomial probability of getting infected. Additional probabilities are used to determine individual's disease course conditioned on the age class of the individual (lethal or transient infection) and the time span being

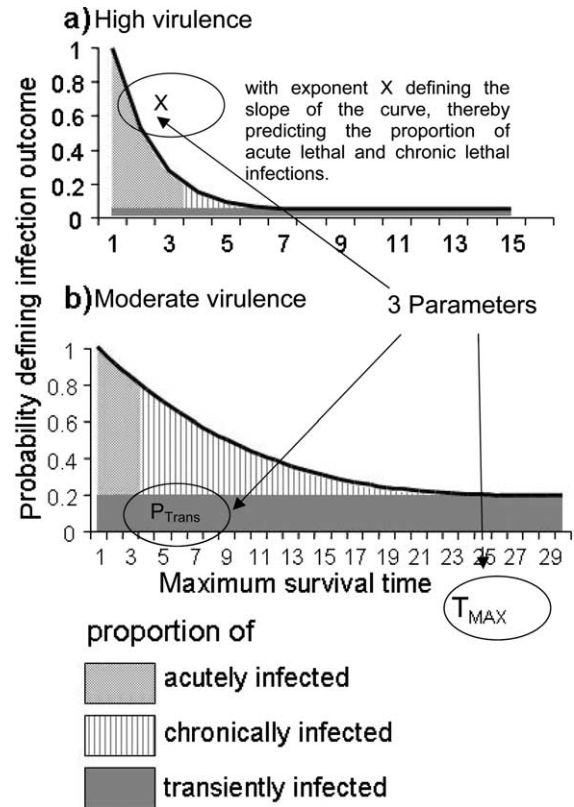


Figure 1. Plot of Eq. 4 (in Supplementary material Appendix 1) defining the disease outcome (survival probability over time) of an individual infection based on a maximum survival time T_{MAX} and an exponent X (lethal infection) or the probability of getting only transiently infected P_{TRANS} , depending on the individual's age class. Different parameter combinations (a) $P_{TRANS} = 0.05$, $T_{MAX} = 15$, $X = 10$ result in about 5% transient infections, 90% acute lethal infections and about 5% chronic lethal infections, or (b) $P_{TRANS} = 0.2$, $T_{MAX} = 29$, $X = 3$ result in about 20% transient infections, 20% acute lethal infections and about 60% chronic lethal infections.

infected (Fig. 1). The latter stochasticity adds individual variations in infectiousness to the model (Supplementary material Appendix 1).

Initialization

We first ran the population model for five years to have a stabilized population structure before one infected animal is released into the wild boar population. The respective ages of the initially released wild boar were drawn from a distribution we obtained from the model after having it run for 100 years. In each run, one infected animal was released in the spring and at a fixed starting point on the landscape to avoid differences in disease persistence time due to random distance from the edge of the simulation area.

Input

The model does not include any external model or data files of driving environmental variables.

Sub-models

The sub-models representing the model's processes are described in the Supplementary material Appendix 1.

Parameters, simulation experiments and analyses

Field data

The demographic parameters of the wild boar model mainly stem from published data from populations in France, Italy, Poland and Germany. The data are based on long-term field studies of up to 10 years of mostly hunted populations (Table A1). For analysis of disease persistence, we fixed parameters for the boar population dynamics according to these references. There are insufficient data available to identify the distribution function for all disease parameters. Therefore, a uniform distribution was assumed for each parameter with upper and lower limits derived either from literature or by estimation. A maximum number of reproducing females per home range (breeding capacity, a correlate of capacitive density) as well as the size of the landscape was also varied (Table A2).

First-order independent variables

We used the following model parameters as independent variables in our statistical analysis testing the persistence hypotheses (Table 1): breeding capacity per home range C_B (hypothesis 1), landscape or habitat-patch size $Size_L$ (hypothesis 2), the effect of the duration of maternal antibodies TS_{MA} (hypothesis 5), and disease outcome (hypothesis 4). The individual disease outcome is modelled in a novel way by three parameters describing the individual responses of the host to infection, namely the probability of getting transiently infected P_{TRANS} , the maximum duration of the infection T_{MAX} and the proportion of acutely and chronically infected individuals given by the exponent X (Fig. 1). Disease transmission follows the effective infection probability within (P_{INF_G}) and between (P_{INF_N}) groups.

Second-order independent variables

From the first-order parameters we also calculated three aggregated model parameters that we used as second-order independent variables: N_{TOT} , being the total initial population size as the product of home range numbers and density, T_{INF} , being the mean infectious period of an infected individual depending on P_{TRANS} , T_{MAX} and X (Eq. A4), the mean-to-variance-ratio VAR_{INF} of the infectious period T_{INF} and S_{INF} , being a measure for how fast the disease spreads through the landscape or the force of infection, depending on the effective infection probability within (P_{INF_G}) and between (P_{INF_N}) groups (Table A2).

Simulation experiments

For all variable values in the model their possible range was divided into 25 equal intervals, and thereof 50 parameter combinations were assigned with a latin hypercube. We ran 20 different hypercubes, resulting in 1000 different parameter combinations. Each parameter combination was run with 120 replicates so that disease persistence could be determined with a precision of 99% with 95% confidence.

The simulations were run for 10 years. To test hypothesis 3 the runs were conducted with and without the presence of prenatally infected piglets PI for 64 selected extreme and intermediate variable combinations, and the effect of adding PI to the disease system compared with a Mann-Whitney U-test (Table A2).

Dependent response variables

We measured persistence counts C_{PERS} , i.e. how often the disease persisted for more than 10 years in the 120 replicates. In addition, we estimated an individual-based measure similar to the analytical R_0 to assess the relationship between the basic reproduction number and disease persistence in our model (Table A2). According to the effective reproduction number R_t (Anderson et al. 2004), which is defined as the number of infections caused by each new case occurring at time t , we introduce \hat{R}_t . Our measure \hat{R}_t is the averaged ratio between already infected and newly infected individuals per time step, i.e. the mean number of new cases emerging from an infected individual per time step, multiplied by the average life-time of an infected individual T_{INF} . With this, we get the average number of secondary infections per life time of a primarily infected individual, measured from the beginning of the outbreak until the end of the outbreak or simulation. Thus, \hat{R}_t is the averaged R_t times T_{INF} and does not separate between early and late epidemic stages. As \hat{R}_t is an estimate from model output, it will change with stochasticity, the landscape structure and hence total population size, and disease parameters. Thus, our measure \hat{R}_t will differ from the traditional analytical R_0 value, which is constant and assuming an infinite and well-mixed fully susceptible population (Cross et al. 2005, 2007, Breban et al. 2007).

Analysis

We distinguish different phases in the development of an outbreak: early extinctions or fade-out, epidemic (i.e. primary outbreak; disease invasion) and endemic situations (i.e. secondary outbreaks to long-term persistence). Early extinctions were those runs in which the outbreak did not manage to infect the initial population before fade-out. In these situations \hat{R}_t is very similar to the analytical R_0 . Runs we termed invasive were those runs that went through the whole landscape of the initial, fully susceptible population, i.e. infecting at least one individual in each cell or home range in the landscape. Persistent runs were invasive outbreaks that in most cases (frequency of persistence $\geq 95\%$) were not self-limiting within a time span of 10 years. We then investigated how the analytical threshold $R_0 > 1$ related to \hat{R}_t under stochastic and heterogeneous conditions, and how \hat{R}_t changed in the different phases of an outbreak.

We analyzed disease persistence in two steps according to the phases of an outbreak: in step 1 we ranked the factors according to their strength to separate early fade-outs from successful invasion using a logistic regression on the outcome of the single runs (procedure glm in R with binomial error distribution and logit link function, R Core Development Team, Table 2). In the following analyses we disregarded parameter combinations producing early-fade

Table 2. Result of the logistic regression of the single runs separating early fade-outs from successful disease invasion. We mark the variables with the strongest influence (estimate of significant variables) in bold. (1) Model based on first-order variables, (2) Model based on second-order variables.

Model parameter	Description	Estimate	SE	z-value	p	Rank
(1) Intercept		−0.7	0.6	−10.9	B0.0001	
P_{INF_G}	Effective infection probability within herd	5.1	0.08	64.5	B0.0001	1
C_B	Breeding capacity per home range (cell)	4.2	0.05	77.0	B0.0001	2
P_{INF_N}	Effective infection probability between herds being the fraction of P_{INF_G}	3.4	0.25	13.4	B0.0001	3
P_{TRANS}	Probability of transient infection (for sub-adults; values for adults and piglets have to be calculated with the formula described in the Digital Appendix)	−0.8	0.05	−15.7	B0.0001	
X	Exponent, giving the proportion of chronic and acute infections	−0.5	0.04	−10.5	B0.0001	
TS_{MA}	Number of weeks, where piglet/ sub-adult is protected by maternal antibodies	−0.4	0.04	−11.7	B0.0001	
T_{MAX}	Maximum survival time (weeks) of lethally infected boars	0.1	0.04	3.0	0.003	
$Size_L$	Landscape size (number of cells). One dimension is standardized with 25 cells	−0.08	0.04	2.0	0.05	
(2) Intercept		−0.46	0.03	−16.1	B0.0001	
S_{INF}	Overall infection probability giving the speed of spread	17.2	0.20	88.3	B0.0001	1
N_{TOT}	Total population size	3.4	0.07	46.3	B0.0001	2
T_{INF}	Mean survival time of an infected individual	1.3	0.1	13.4	B0.0001	3
VAR_{INF}	Proportion of long-term virus shedders	−0.6	0.2	−3.6	B0.001	

out, hence we considered only parameter combinations which were invasive or persistent. In step 2 we separated successful invasion from persistence. We inspected the association of C_{PERS} and \hat{R}_t using Pearson's product moment correlation coefficient and calculated Kendall's tau partial rank correlation for a first inspection of the association between independent variables and output measures (Fig. A2A, A2B). We then evaluated the structural relationship between the first- and second-order independent variables and C_{PERS} using a generalized additive model (GAM, procedure gam with binomial error distribution, Wood 2006). We finally calculated generalized linear models (GLM) for C_{PERS} and \hat{R}_t based on the first- and second-order independent variables (procedures glm with binomial error distribution for C_{PERS} and with Poisson error distribution for \hat{R}_t , R 2.5.1, Table 3). All independent variables were standardized between 0 and 1 to assess their relative contribution to invasion and persistence. We then rank the variables based on their estimate in the full model, i.e. all variables or hypotheses, respectively, were simultaneously confronted and ranked.

Results

We found strong synchronization in outbreak peaks associated with the strong annual reproductive pattern (Fig. 2). There obviously is a critical phase for the virus after the primary outbreak, i.e. when most susceptibles are either infected, dead or immune, and in winter as well when there are no new susceptibles.

In step 1 we separated early fade-out from invasion. In agreement with epidemic theory of disease spread in naïve host populations, the density (C_B) and force of the infection within groups (P_{INF_G}) were the most important variables when predicting invasion. Then variables describing the individual disease course follow (Table 2(1)). The landscape size was less important. In contrast, the aggregated model based on second-order variables, although yielding qualitatively similar results, was not able to disentangle the effect of global population size and local group density (Table 2(2)).

In step 2 we separated invasion from long-term persistence. Since all first- and second-order variables

Table 3. Result of the generalized linear models separating successful invasion only from long-term disease persistence in for invasive runs. The most important parameters (rank, description see Table 2) are marked in bold. (1) Model based on first-order variables; (2) model based on second-order variables.

(A) Persistence counts C_{PERS}						(B) Counts of secondary infections \hat{R}_t					
	Model parameter	Estimate	SE	z-value	p	Rank	Estimate	SE	z-value	p	Rank
(1)	Intercept	−4.9	0.05	−109.5	B0.001		0.7	0.1	6.7	B0.001	
	T_{MAX}	2.7	0.03	84.7.6	B0.001	1	0.01	0.07	0.2	0.8	
	P_{TRANS}	2.3	0.03	71.5	B0.001	2	−0.6	0.08	−7.7	B0.001	2
	X	−1.9	0.03	−62.0	B0.001	3	−0.008	0.07	−0.1	0.9	
	$Size_L$	1.6	0.03	53.9	B0.001	4	−0.5	0.07	−7.1	B0.001	3
	C_B	1.3	0.03	46.2	B0.001	5	0.4	0.07	5.6	B0.001	4
	P_{INF_N}	−0.9	0.07	−14.1	B0.001	6	0.7	0.14	4.7	B0.001	1
	P_{INF_G}	0.14	0.03	4.3	B0.001		0.3	0.1	3.6	B0.001	5
	TS_{MA}	−0.01	0.02	−0.7	0.5		−0.2	0.06	2.4	0.016	
(2)	Intercept	−1.0	0.3	−39.38	B0.001		0.5	0.05	10.6	B0.001	
	VAR_{INF}	−262.8	4.4	−60.4	B0.001	1	0.3	0.5	0.7	0.5	
	N_{TOT}	2.1	0.03	67.9	B0.001		−0.2	0.1	−2.4	0.02	
	T_{INF}	−1.2	0.06	−19.6	B0.001		0.7	0.1	5.1	B0.001	2
	S_{INF}	−0.6	0.06	−10.1	B0.001		1.0	0.1	7.8	B0.001	1

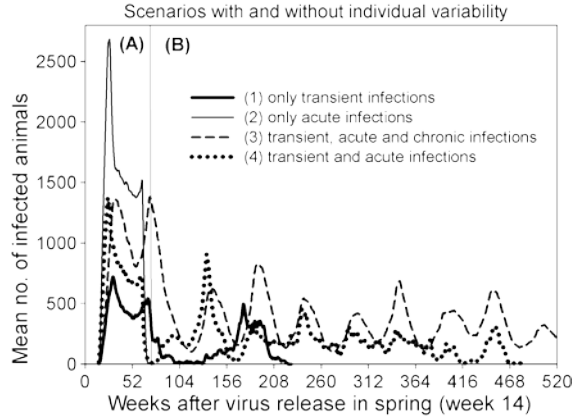


Figure 2. Mean number (10 repetitions) of transiently and lethally infected individuals with different levels of individual variability in the disease outcome. (A) shows the phase of the disease invasion, and (B) the phase of endemic cycling. The graphs show that there is a critical phase for the virus after the primary outbreak; thereafter synchronized yearly cycling of outbreaks occurs. Only when adding variability to the disease outcome, i.e. having transient, acute and chronic infections occurring at the same time (example 3), is the disease likely to persist. Disease parameters: (1) $P_{INF_G}=0.1$, $P_{TRANS}=1$; (2) $P_{INF_G}=0.5$, $P_{TRANS}=0$, $X=1$, $T_{MAX}=5$; (3) $P_{INF_G}=0.1$, $P_{TRANS}=0.7$, $X=4$, $T_{MAX}=30$; (4) $P_{INF_G}=0.5$, $P_{TRANS}=0.7$, $X=1$, $T_{MAX}=5$. (Fixed: $C_B=5$, $Size_L=1250$, $P_{INF_N}=10$).

showed monotonic response in the GAMs, they were included as linear in the GLMs (Fig. A3, results not shown for the second-order variables). We could reject two hypotheses: An additional proportion of chronically infected individuals through prenatally infected offspring had no significant effect on persistence C_{PERS} (hypothesis 3, Mann-Whitney-U-test, $p=0.92$). The effect of a low level of maternal antibodies TS_{MA} was also negligible, and therefore the associated hypothesis was rejected (hypothesis 5, Table 3A1). The infection probabilities within and between groups only played a minor role; this is because we only considered the invasive parameter combinations, which were characterized by a sufficiently large infection probability.

Disease traits (hypothesis 4), landscape extent (hypothesis 2) and host density (hypothesis 1) were the main contributors to disease persistence (Table 3). In detail, cycling of the disease was favored by a high probability of transient infections P_{TRANS} as well as many chronic long-term shedders (low exponent values X and long infection times T_{MAX} ; Table 3A1). We found a significant, but low correlation between P_{TRANS} and the mean number of animals in the simulations (Pearson's $r=0.33$, $p<0.001$). Thus, the mechanistic basis of P_{TRANS} is not solely related to multiplying the number of animals in the system.

Large landscapes or habitat patches, respectively ($Size_L$; Table 3A1), also contributed to persistence, as new infections in the next generation (susceptible offspring in spring) could start while the primary outbreak was still in the invasive phase. Together with a high wild boar density (C_B ; Table 3A1) leading to high numbers of susceptible piglets, this contributed to disease persistence. Thus, when disease characteristics allow stochastic fluctuations in the outcome of the epidemic, this epidemic is more likely to

persist in larger populations. These results are also reflected in the second-order model (Table 3A2).

To check whether \hat{R}_t contributes to an understanding of persistence, we stratified simulation runs according to the three different outbreak categories early fade-out, invasion, and long-term persistence (Fig. 3A). We found higher \hat{R}_t values in invasion and persistence cases than in fade-out situations, although the highest \hat{R}_t values show low persistence (Fig. 3B). The analytical threshold for invasion, $R_0>1$ was also reflected by our measure \hat{R}_t as separating between fade-out and invasion or persistence. However, the separation was not as clear-cut as in analytical theory.

The correlation between the two response measures C_{PERS} and \hat{R}_t was significant, but low and negative (Pearson's $r=-0.21$, $p<0.001$), indicating that these two measures act in opposing directions. Also, highest \hat{R}_t values had the lowest probability of persistence (Fig. 3B). If persistence was just an effect of increasing \hat{R}_t , which entails transmissibility and the infectious period, then we would expect S_{INF} and T_{INF} (or their first-order equivalents) to also be the main factors contributing to persistence. We

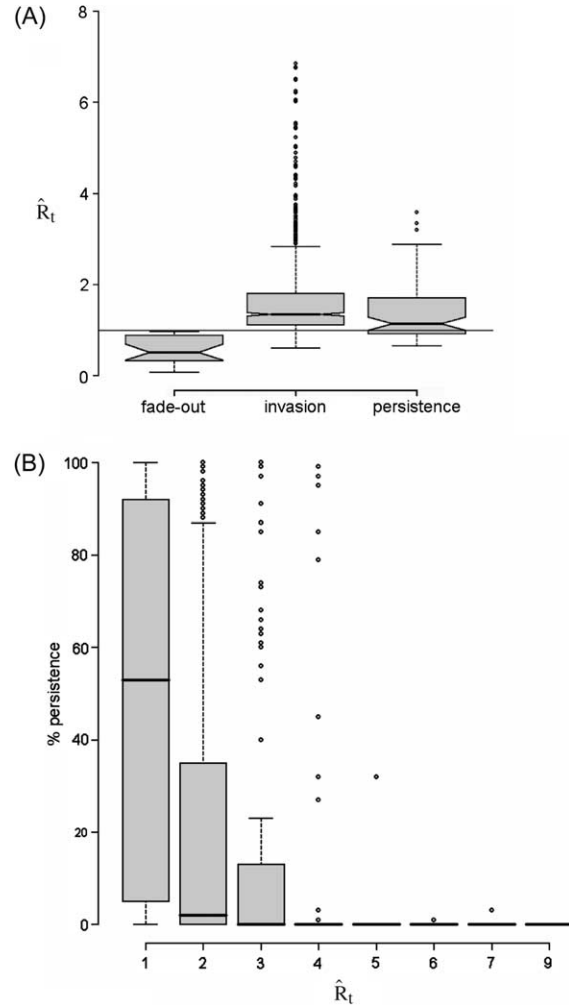


Figure 3. Boxplots of the medians for (A) \hat{R}_t in the different phases of an outbreak. The horizontal line shows the analytical threshold for successful disease invasion or extinction. (B) different probabilities of disease persistence for \hat{R}_t for the 2 categories of (A): invasion and persistence.

therefore compared the processes that led to a high \hat{R}_t with those that yielded persistence (Table 3B).

The second-order model, where the first-order variables were aggregated in the same manner as in standard epidemiological models, reflected the coarse picture (Table 3B2). In agreement with theory, \hat{R}_t was mainly dependent on transmission S_{INF} and the mean infectious period T_{INF} . The decisive difference for uncoupling \hat{R}_t processes from persistence processes is the effect of the variability in the infectious period VAR_{INF} on persistence (Table 3A2). On the other hand, the factor P_{INF_N} , which represents how fast the disease can spread through the landscape, was most important for explaining \hat{R}_t (rank 1) but was not important when explaining persistence (rank 6; Table 1,3).

Discussion

To disentangle factors contributing to the persistence of wildlife diseases that persist without a reservoir host, we developed an individual-based spatially explicit model. We analyzed this simulation model using descriptive models and found that the features of host and disease that contributed most to long-term persistence of the disease were characterized by high individual variability, i.e. host populations with mostly transient, i.e. surviving individuals requiring immunity, and some chronically, long-term infected animals. We systematically analyzed the model at two different aggregation levels (first- and second-order model parameters) and used two different ‘currencies’ for characterizing epidemics: probability of persistence over 10 years, and the individual-based equivalent of the basic reproductive number R_0 , \hat{R}_t , which was the average number of secondary infections per life time of a primarily infected individual. In the following we first discuss the modeling approach, compare our understanding of disease invasion and persistence at the population and process level, discuss lessons for disease persistence, and finally discuss management implications of our results.

The modeling approach

While many diseases have complex dynamics, they also depend on the dynamics of an animal community. To understand, predict, and control diseases, we must go beyond simple analytically formulated models that provide an insight into general disease dynamics, but may be of limited utility to evaluate the role of ecological details critical in specific host-parasite systems. There are two main challenges in the modeling approach, the first being the development of more complex models accounting for individual variability, nonlinearity, stochasticity and space, and the second challenge being the handling of the sparse data that exist on diseases in wildlife populations (Hastings et al. 2005). We tackled these challenges by developing a moderately complex individual-based, spatially explicit model combining ecological traits of the hosts as well as characteristics of a viral disease. We incorporated a range of disease persistence hypotheses because lack of conclusive data should not be a barrier to exploratory modeling of the system, as it can lead to general conclusions about critical

persistence factors (Smith et al. 2001, Caley and Hone 2004).

The novelty of our modeling approach is the variability of the disease outcome on the individual level. The parameters characterizing the disease course in the individual were based on observations made in experiments. Usually, either case mortality or expected lifespan of infected hosts is used to study how pathogens harm their host (Day 2002), and often the host population is modeled homogeneously in its resistance characteristics (Gandon et al. 2002). In our approach, the disease outcome could be modulated with respect to both factors. Epidemiologists have also pointed out that the force of infection is a dynamic variable that depends on the number of infected hosts in the population (Gandon et al. 2002). As we modeled disease transmission at the level where the process occurs – individual-to-individual transmission – becoming infected was no longer based on a constant transmission rate (McCallum et al. 2001) but changed over the course of time depending on the number of infected individuals in the vicinity, thereby capturing the essential characteristics of the system (Turner et al. 2003).

Disease invasion and persistence – comparing population and process level

Normally, population-level analyses use average quantities to describe heterogeneous systems. A prominent example, central to the current understanding of epidemic spread, is the basic reproductive number R_0 . However, population estimates of R_0 can obscure considerable individual variation in infectiousness (Keeling and Grenfell 2000, Lloyd-Smith et al. 2005b). Cross et al. (2005) have analyzed the behavior of a measure that is equivalent to R_0 in an individual-based context and found that $R_0 > 1$ is not a sufficient condition for invasion success, and that invasion success is determined by many other factors, such as the population turnover rate (Cross et al. 2007, Breban et al. 2007).

Our individual-based measure \hat{R}_t predicted disease invasion fairly well, but completely failed with further separation of invasion from long-term persistence (Fig. 3A). The correlation between the measure of persistence (C_{PERS}) and the measure of transmission intensity (\hat{R}_t) is minimal, i.e. an increase in \hat{R}_t does not coincide with an increase in C_{PERS} .

However, there are hypotheses stating that factors which increase R_0 (transmission probability and mean infectious period) will usually increase disease persistence. Therefore, we compared the processes that led to high persistence C_{PERS} or high \hat{R}_t , respectively. We first aggregated our disease variables into variables similar to those in mean field approaches and found that the disease features ‘mean infectious period’ (T_{INF}) and ‘mean infection intensity’ (S_{INF}) are no longer considered to be important in the explanation of persistence. The variability in the infection period (VAR_{INF}) was the most important epidemiological factor for persistence (Table 3A2).

Our results revealed limitations of R_0 as an indicator characterizing endemic situations. R_0 is not meant to reflect variability but mean properties of a disease. The condition $R_0 > 1$ may apparently be applicable to explain disease

invasion and partly persistence, but it is concealing underlying processes. The reason is that R_0 measure is not prepared to distinguish between non-persisting invasive outbreaks from those that actually persisted.

We detected this by considering how our individual-based equivalent of R_0 , \hat{R}_t , depended on different processes (Table 3B1). We found two reasons why \hat{R}_t did not explain persistence: (1) the variability in infection period is simultaneously shortened (P_{TRANS} negatively associated with \hat{R}_t) and prolonged by a chronic course (T_{MAX} positively associated), hence the main aspects of disease variability counteract when influencing \hat{R}_t . (2) It is especially the transmission within and between the home range cells that is the important difference for maximizing \hat{R}_t in comparison to variables maximizing persistence (Table 3B1), i.e. spreading as fast as possible through the landscape. This mechanism is of minor importance when explaining persistence (given already successful invasion).

Indeed, persistence is a more local process; local or 'dormant' clusters with long-term shedders that spread only slowly to neighboring cells in a structured population build temporal bridges until the next reproductive cycle based on immune survivors fills up the susceptible pool. This does not work if the whole population is infected too quickly, making \hat{R}_t a counterproductive measure to explain persistence. This linking of the process-based view on disease dynamics to the usual population level view could be used as a more general framework for modeling epidemics in cases where we can assume that individual's effects on disease dynamics and the interaction among intervening factors of a very different nature are important.

Lessons for disease persistence

In CSF literature, the increase in endemic situations has been hypothesized to become more frequent in recent years because of the increasing size and density of the wild boar population and the involvement of viral strains of moderate to low virulence. It is difficult to empirically or descriptively disentangle the relative contribution of both factors. Our study revealed the striking role of individual variability in the disease outcome. The degree of this variability determines not only the percentage of individuals that recover to immunity (i.e. transient infections), but also the amount of long-lasting infections (i.e. chronic courses). Such individuals can sustain the outbreak by providing infection carriers through periods where the incidence of infection is low due to reduced host reproduction. On the other hand, a high proportion of transiently infected animals guarantees the survival of most wild boars and therefore a constant supply of susceptible offspring which stabilizes the infection chain during reproductive periods (Lloyd-Smith et al. 2005a) and which keeps the host population above the critical community size (Conlan and Grenfell 2007).

It is known from different laboratory experiments that the outcome of an infection can vary widely with respect to the relative frequency of individual outcomes (i.e. transient, or lethal (either acute or chronic)). This knowledge motivated the epidemiological disease concept in our model. Now our findings have disclosed the prime

importance of the individual variation in infectiousness for the persistence of a given epidemic in the wild.

This finding is underpinned by recent genotyping analyses showing a switch from highly virulent CSF viruses of group 1 to moderately virulent viruses of group 2 since the 1970s (Paton et al. 2000). The virulence evolution hypothesis (VEH) claims that strains of intermediate virulence are the outcome of natural selection to balance between the need of the virus to reproduce and the costs of harming the host (Gandon et al. 2002, Zimmer 2003, Day and Proulx 2004, but see Iwasa et al. 2005). From other species and their diseases it is hypothesized that a non-pathogenic strain of rabbit haemorrhagic disease virus seems to protect rabbits from virulent outbreaks (White et al. 2001).

So far, mainly the acute course of the CSF disease has been described in the field (Dahle and Liess 1992). Further, there is first evidence for transient infections from serological studies in wild boar (Rossi et al. 2005a). Although convincingly demonstrated in animal experiments, evidence is missing to date for chronically infected individuals in infected wild boar populations. This can be due to the difficulty in detecting them in the field, as they do not necessarily show the typical clinical signs. However, new data from intensive field surveys of infected wild boar populations hint at the existence of long-lasting infections in the wild (Sophie Rossi pers. comm.). Our model results suggest that chronically infected wild boars must be present in populations with long-lasting epidemics, and therefore there is an urgent need to test this hypothesis in the field.

The second facet of our findings, i.e. the significance of a large population size via large landscapes and high density, relates disease persistence to the mechanism included in more general epidemic models. Our separated analysis of invasion and persistence, however, figured the conceptual difference: When the infection is first introduced to the population the virus has sufficient access to susceptible animals. This situation is well described by the standard epidemic population model (Anderson and May 1979, Barlow 1996). Hence, the invasibility of the host population by the infection is expected to be determined according to the predictions of standard epidemiological models: multiplication of an infection in a completely susceptible population hinges on a threshold density which depends on the basic reproductive number of the infectious disease, i.e. R_0 . Therefore, we would expect an outbreak successfully passing through the host population depending on the force of infection and host density (Grenfell and Dobson 1995), a pattern that our model has reproduced. Here, it should be noticed that our analysis highlighted the size of the habitat patch to be completely of non-interest when enquiring about successful invasion. This indeed appears reasonable when we recall that epidemic invasibility is a threshold phenomenon. Hildenbrandt et al. (2006) found the same in their analysis of the establishment and persistence of small populations: establishment is not affected by habitat capacity, because the founder population, which usually is spatially restricted, is affected by local factors, not the capacity of the entire habitat.

The situation is different, however, when the epidemic has passed through the whole landscape of susceptible hosts and is still persistent. As known from general infectious

diseases modeled explicitly in space (Mollison and Levin 1995), particular features of the infection like long-term shedders are needed to link the last infections of the prime wave front to newly available hosts that could continue the chain of infection. Our results found such disease attributes promoting long-term persistence: they cause long-time surviving virus shedders and preserve reproduction via many transient infectious courses.

Habitat patch size was, in contrast to invasion success, highly important when explaining persistence. The effect of the habitat patch size parameter is related to bridging the time gap between the first and secondary outbreak in such that the invasive front is still traveling through the fully susceptible population while the surviving individuals behind the wave front provide new susceptibles that are infected via the first outbreak. Additionally, a large landscape with many individuals is likely to increase the probability of observing chance events such as long-time surviving carriers, i.e. a population of hosts occupying large spatial extensions increases the likelihood of disease persistence simply because more clusters of infection can remain behind. This is not related to a direct contact chain of infection. Thus, population-level non-spatial models could fail both on the correct evaluation of the effects of habitat patch size vs contact density as well as on the identification of prominent causes for persistence of the infection.

Implications for disease management

Vaccination and hunting are presently considered as eradication measures for persistent diseases in wildlife. From our model results we learned that the CSF seems to persist through the youngest age class, i.e. the newly born susceptibles, because after the primary invasive outbreak, most adults will be either immune or dead. But the youngest age class is very difficult to vaccinate, because they do not take up the baits until the age of 4–5 months (Volker Kaden and Sophie Rossi pers. comm.). Thus, vaccination could also have negative effects by artificially keeping alive the reproductive adults that would otherwise have died from the disease, producing susceptible offspring that cannot be vaccinated. The same can be concluded for hunting. Reducing the susceptible age class through hunting as an eradication strategy logically follows from the paradigm of a threshold density for the invasion and a critical community size for the persistence of a disease (Caley and Ramsey 2001), but can have reverse effects if we think of a density-dependent population regulation (Choisy and Rohani 2006) leading to many susceptible piglets in the following spring or disturbance effects (Vicente et al. 2007).

In a recent work using CSF as an example, it is shown that the presence of age-dependent heterogeneity in the transmission rate may produce the counter-intuitive result that disease prevalence increases over a range of intermediate levels of culling (Bolzoni et al. 2007). Similarly, our results indicate that early intervention by vaccination might be counter-productive: immune animals by vaccination mimic the worse outcome of a transient infection. We thus conclude that the effects of hunting and vaccination might strongly interact with ecological processes of the complex disease-host system. Detailed assessment using appropriate

models is required to deduce the optimum control scheme—also in relation to costs.

Acknowledgements – We thank Jürgen Teuffert, Christoph Stau-bach, Klaus Depner, Volker Kaden, Sophie Rossi and Jane Uhd Jepsen for fruitful discussions and valuable comments. Thanks to Paul Ronning for correction of the English language and Michael Müller for assistance in programming. SKS and NF were supported by a Marie Curie Host Fellowship provided by the European Commission (SKS: MEIF-CT-2006-39985, NK: HPMD-CT-2001-00109).

References

- Anderson, R. M. and May, R. M. 1979. Population biology of infectious diseases: part I. – *Nature* 280: 361–367.
- Anderson, R. M. and May, R. M. 1991. Infectious diseases of humans: dynamics and control. – Oxford Univ. Press.
- Anderson, R. M. et al. 2004. Epidemiology, transmission dynamics and control of SARS: the 2002–2003 epidemic. – *Philos. Trans. R. Soc. Lond. B* 359: 1091–1105.
- Artois, M. et al. 2002. Classical swine fever (hog cholera) in wild boar in Europe. – *Rev. Sci. Tech. OIE* 21: 287–303.
- Barlow, N. D. 1996. The ecology of wildlife disease control: simple models revisited. – *J. Appl. Ecol.* 33: 303–314.
- Bieber, C. and Ruf, T. 2005. Population dynamics in wild boar *Sus scrofa*: ecology, elasticity of growth rate and implications for the management of pulsed resource consumers. – *J. Appl. Ecol.* 42: 1203–1213.
- Bolzoni, L. et al. 2007. Transmission heterogeneity and control strategies for infectious disease emergence. – *PLOS One* 2: e747.
- Breban, R. et al. 2007. Theory versus data: how to calculate R_0 . – *PLOS One* 2: e282.
- Caley, P. and Ramsey, D. 2001. Estimating disease transmission in wildlife, with emphasis on leptospirosis and bovine tuberculosis in possums, and effects of fertility control. – *J. Appl. Ecol.* 38: 1362–1370.
- Caley, P. and Hone, J. 2004. Disease transmission between and within species, and the implications for disease control. – *J. Appl. Ecol.* 41: 94–104.
- Caley, P. and Hone, J. 2006. Assessing the host disease status of wildlife and the implications for disease control: *Mycobacterium bovis* infection in feral ferrets. – *J. Appl. Ecol.* 42: 708–719.
- Choisy, M. and Rohani, P. 2006. Harvesting can increase severity of wildlife disease epidemics. – *Proc. R. Soc. Lond. B* 273: 2025–2034.
- Conlan, A. J. K. and Grenfell, B. T. 2007. Seasonality and the persistence and invasion of measles. – *Proc. R. Soc. Lond. B* 274: 1133–1141.
- Cross, P. C. et al. 2005. Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. – *Ecol. Lett.* 8: 587–595.
- Cross, P. C. et al. 2007. Utility of R_0 as a predictor of disease invasion in structured populations. – *J. R. Soc. Interface* 4: 315–324.
- Dahle, J. and Liess, B. 1992. A review on classical swine fever infections in pigs: epizootiology, clinical disease, and pathology. – *Compar. Immunol. Microbiol. Infectious Dis.* 15: 203–211.
- Dardallion, M. 1988. Wild boar social groupings and their seasonal changes in the Camargue, southern France. – *Z. Säugetierkunde* 53: 22–30.

- Day, T. 2002. On the evolution of virulence and the relationship between various measures of mortality. – *Proc. R. Soc. Lond. B* 269: 1317–1323.
- Day, T. and Proulx, S. R. 2004. A general theory for the evolutionary dynamics of virulence. – *Am. Nat.* 163: E40–E63.
- Depner, K. R. et al. 2000. Transient classical swine fever virus infection in wild boar piglets partially protected by maternal antibodies. – *Dtsch Tierarztl. Wochenschr.* 107: 66–68.
- Eisinger, D. and Thulke, H.-H. 2008. Spatial pattern formation facilitates eradication of infectious disease. – *J. Appl. Ecol.* 45: 415–423.
- Fernandez, N. et al. 2006. Viability and risk assessment in species restoration. Planning reintroductions for the wild boar, a potential disease reservoir. – *Ecol. Soc.* 11: 1–6.
- Fritzemeier, J. et al. 2000. Epidemiology of classical swine fever in Germany in the 1990s. – *Vet. Microbiol.* 77: 29–41.
- Galvani, A. P. and May, R. M. 2005. Dimensions of super-spreading. – *Nature* 438: 293–295.
- Gandon, S. et al. 2002. The evolution of parasite virulence, superinfection, and host resistance. – *Am. Nat.* 159: 658–669.
- Gewin, V. 2004. Virtual plagues get real. – *Nature* 427: 774–775.
- Grenfell, B. T. and Dobson, A. P. 1995. Ecology of infectious diseases in natural populations. – Cambridge Univ. Press.
- Grimm, V. and Railsback, S. F. 2005. Individual-based modeling and ecology. – Princeton Univ. Press.
- Grimm, V. et al. 2005. Importance of buffer mechanisms for population viability analysis. – *Conserv. Biol.* 19: 578–580.
- Grimm, V. et al. 2006. A standard protocol for describing individual-based and agent-based models. – *Ecol. Modell.* 198: 115–126.
- Guberti, V. et al. 1998. Estimate the threshold abundance for the persistence of the classical swine fever in the wild boar population of the eastern Sardinia. Report on measures to control classical swine fever in European wild boar. – Commission of the European Communities, Directorate General VI for Agric., pp. 54–61.
- Hansen, F. et al. 2004. Processes leading to a spatial aggregation of *Echinococcus multilocularis* in its natural intermediate host *Microtus arvalis*. – *Int. J. Parasitol.* 34: 37–44.
- Hastings, A. et al. 2005. Quantitative bioscience for the 21st century. – *BioScience* 55: 511–517.
- Hagenaars, T. J. et al. 2004. Spatial heterogeneity and the persistence of infectious diseases. – *J. Theor. Biol.* 229: 349–359.
- Hildenbrandt, H. et al. 2006. How to detect and visualize extinction thresholds for structured PVA models. – *Ecol. Modell.* 191: 545–550.
- Hosseini, P. R. et al. 2004. Seasonality and wildlife disease: how seasonal birth, aggregation and variation in immunity affect the dynamics of *Mycoplasma gallisepticum* in house finches. – *Proc. R. Soc. Lond. B* 271: 2569–2577.
- Iwasa, Y. et al. 2005. Virus evolution within patients increases pathogenicity. – *J. Theor. Biol.* 232: 17–26.
- Keeling, M. J. and Grenfell, B. T. 2000. Individual-based perspectives on R-0. – *J. Theor. Biol.* 203: 51–61.
- Kelly, J. K. et al. 2003. Linking dynamical and population genetic models of persistent viral infection. – *Am. Nat.* 162: 14–28.
- Kern, B. et al. 1999. Incidence of classical swine fever (CSF) in wild boar in a densely populated area indicating CSF virus persistence as a mechanism for virus perpetuation. – *J. Vet. Med. B* 46: 63–67.
- Kramer-Schadt, S. et al. 2007. A review of potential ecological and epidemiological factors affecting the persistence of classical swine fever in wild boar populations. – *Mamm. Rev.* 37: 1–20.
- Leaper, R. et al. 1999. The feasibility of reintroducing wild boar (*Sus scrofa*) to Scotland. – *Mamm. Rev.* 29: 239–259.
- Lloyd-Smith, J. O. et al. 2005a. Should we expect population thresholds for wildlife disease? – *Trends Ecol. Evol.* 20: 511–519.
- Lloyd-Smith, J. O. et al. 2005b. Superspreading and the effect of individual variation on disease emergence. – *Nature* 438: 355–359.
- Mackinnon, M. J. and Read, A. F. 2004. Virulence in malaria: an evolutionary viewpoint. – *Philos. Trans. R. Soc. Lond. B* 359: 965–986.
- McCallum, H. et al. 2001. How should pathogen transmission be modelled? – *Trends Ecol. Evol.* 16: 295–300.
- Meyers, G. and Thiel, H.-J. 1996. Molecular characterization of pestiviruses. – *Adv. Virus Res.* 47: 53–118.
- Mollema, L. et al. 2005. Prolonged persistence of bovine herpesvirus in small cattle herds: a model-based analysis. – *Epidemiol. Infect.* 133: 137–148.
- Mollison, D. and Levin, S. A. 1995. Spatial dynamics of parasitism. – In: Grenfell, B. T. and Dobson, A. P. (eds), Ecology of infectious diseases in natural populations. Cambridge Univ. Press, pp. 384–398.
- Orive, M. E. et al. 2005. Viral infection in internally structured hosts. I. Conditions for persistent infection. – *J. Theor. Biol.* 232: 453–466.
- Paton, D. J. et al. 2000. Genetic typing of classical swine fever virus. – *Vet. Microbiol.* 73: 137–157.
- Rossi, S. et al. 2005a. Long-term monitoring of classical swine fever in wild boar (*Sus scrofa* sp.) using serological data. – *Vet. Res.* 36: 27–42.
- Rossi, S. et al. 2005b. Incidence and persistence of classical swine fever in free-ranging wild boar (*Sus scrofa*). – *Epidemiol. Infect.* 133: 559–568.
- Smith, G. C. et al. 2001. A model of bovine tuberculosis in the badger *Meles meles*: the inclusion of cattle and the use of a live test. – *J. Appl. Ecol.* 38: 520–535.
- Turner, J. et al. 2003. Modelling pathogen transmission: the interrelationship between local and global approaches. – *Proc. R. Soc. Lond. B* 270: 105–112.
- Vicente, J. et al. 2007. Social organisation and movement influence the incidence of bovine tuberculosis in an undisturbed high-density badger *Meles meles* population. – *J. Anim. Ecol.* 76: 348–360.
- White, P. J. et al. 2001. The emergence of rabbit haemorrhagic disease virus: will a non-pathogenic strain protect the UK? – *Philos. Trans. R. Soc. Lond. B* 356: 1087–1095.
- Wood, S. N. 2006. Generalized additive models: an introduction with R. – Taylor and Francis CRC Press.
- Zimmer, C. 2003. Taming pathogens: an elegant idea, but does it work? – *Science* 300: 1362–1364.

Supplementary material (available online as O16582 at www.oikos.ekol.lu.se/appendix). Appendix 1.